REMARKS

Claims 1, 2-6, 8-23, and 25-35 were previously pending.

Claims 1 and 28 are amended.

Claims 2, 8, 20-23, and 25-27 are canceled.

New claims 36-39 are added.

Claims 1, 3-6, 9-19, and 28-39 are pending.

Rejections under 35 U.S.C. §103(a)

Claims 1-6, 8-23, and 25-35 were rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,188,976 to Ramaswamy et al., filed Oct. 23, 1998 ("Ramaswamy") in view of U.S. Patent No. 6,317,707 to Bangalore et al., filed Dec. 7, 1998 ("Bangalore").

Without regard for their rejection, Applicant requests that claims 2, 8, 20-23, and 25-27 be canceled, and that the cancellation be made without prejudice so that these canceled claims may be pursued, at the Applicant's discretion, in a continuation application. Current rejection of these claims is moot because of their cancellation.

Claims 1 and 28 have been amended for clarity, to "more particularly point out and distinctly claim the subject matter." The amendments of claims 1 and 28 are not meant to narrow the scope of these claims but merely to increase their clarity.

Claim 1

Claim 1, as amended, defines a method of using a tuning set of information to jointly optimize the performance and size of a language model, including:

> segmenting at least a subset of a received textual corpus into segments by clustering every N-items of the received corpus into a training unit, wherein resultant training units are separated by gaps;
>
> creating the tuning set from application-specific information;
>
> (a) training a seed model via the tuning set;
>
> (b) calculating a similarity within a sequence of the training units on either side of each of the gaps;
>
> (c) selecting segment boundaries that maximize intra-segment similarity and inter-segment disparity;
>
> (d) calculating a perplexity value for each segment based on a comparison with the seed model;
>
> (e) selecting some of the segments based on their respective perplexity values to augment the tuning set;
>
> iteratively refining the tuning set and the seed model by repeating parts (a) through (e) until a threshold; and
>
> refining the language model based on the seed model.

Applicant's claim 1 performs an optimization of a language model's performance and size parameters, jointly. Further, Applicant's subject matter can decrease memory usage in a computing device while increasing performance and application-specificity of the language model.

On the other hand, Ramaswamy and Bangalore, either alone or in combination, do not teach or suggest one or more of the elements of Applicant's claim 1.

As just one example, Ramaswamy and Bangalore, either alone or in combination, do not teach or suggest iteratively refining a seed model by iteratively refining a tuning set, which in turn refines the seed model, etc.

Thus, Applicant respectfully submits that amended claim 1 is not taught or suggested by Ramaswamy and Bangalore, either alone or in combination. The combination of Ramaswamy in view of Bangalore fails, and claim 1 should be patentable over the combination of Ramaswamy in view of Bangalore.

Claims 3-6, and 9-19

For at least the reasons set forth above with respect to claim 1, Applicant submits that claims 3-6, and 9-19 are patentable over Ramaswamy in view of Bangalore. Dependent claims contain the language of the claims from which they depend. Claims 3-6, and 9-19 depend directly or indirectly from claim 1. Claim 1 is allowable, therefore, claims 3-6, and 9-19 are also allowable.

Claim 28

Claim 28, as amended, defines a modeling agent, including:

a controller, to receive invocation requests to develop a language model from a corpus; and

a data structure generator, responsive to the controller, to:

develop a seed model from a tuning set of information;

segment at least a subset of a received corpus, wherein the segments of the received corpus are a clustering of every N items of the received corpus into a training unit and the training units are separated by gaps;

calculate the similarity within a sequence of training units on either side of each of the gaps;

select segment boundaries that improve intra-segment similarity and inter-segment disparity;

calculate a perplexity value for each segment;

refine the seed model with one or more segments of the received corpus based, at least in part, on the calculated perplexity values;

iteratively refine the tuning set with segments ranked by the seed model and in turn iteratively update the seed model via the refined tuning set;

filter the received corpus via the seed model to find low-perplexity segments; and

train the language model via the low-perplexity segments.

As with claim 1 above, Ramaswamy and Bangalore, either alone or in combination, do not teach or suggest one or more of the elements of Applicant's claim 28.

For example, Ramaswamy and Bangalore, either alone or in combination, do not teach or suggest an element to "iteratively refine the tuning set with segments ranked by the seed model and in turn iteratively update the seed model via the refined tuning set..."

Thus, Applicant respectfully submits that amended claim 28 is not taught or suggested by Ramaswamy and Bangalore, either alone or in combination, and that the combination of Ramaswamy in view of Bangalore fails. Applicant accordingly suggests that claim 28 is patentable over the combination of Ramaswamy in view of Bangalore.

Claims 29-35

For at least the reasons set forth above with respect to claim 28, Applicant submits that claims 29-35 are patentable over Ramaswamy in view of Bangalore. Dependent claims contain the language of the claims from which they depend. Claims 29-35 depend directly or indirectly from claim 28. Claim 28 is allowable, therefore, claims 29-35 are also allowable.

New Claims 36-39

Claims 36-39 define a method of jointly optimizing the performance and size of a language model in which iteratively training a seed model includes updating the seed model according to the tuning sample, ranking each of multiple segments from a corpora according to a perplexity comparison with the seed model, selecting some of the multiple segments that possess a low perplexity, and augmenting the tuning sample with the selected segments, then repeating until a threshold is reached.

CONCLUSION

Applicant respectfully suggests that claims 1, 3-6, 9-19, and 28-39 are in condition for allowance. Applicant requests allowance of claims 1, 3-6, 9-19, and 28-39 and issuance of the subject application. Should any matter in this case remain unresolved, the undersigned attorney respectfully requests a telephone conference with the Examiner to resolve any such outstanding matter.

Respectfully Submitted,

Date: 8-17-05

By: 

Lee & Hayes PLLC
Mark C. Farrell
Reg. No. 45,988
(509) 324-9256